Aalto University
School of Science
Degree Programme in Industrial Engineering and Management

Jonne Viitanen

Estimation of Risk Premium for German Government Bonds Using Regression Based Affine Model

Master's Thesis

Espoo, May 26, 2015

Supervisor: Professor Hannele Wallenius

Instructors: Ruth Kaila, D.Sc. (Tech.) and Lauri Hälikkä, M.Sc.

AALTO UNIVERSITY SCHOOL OF SCIENCE	ABSTRACT OF THE
Degree Programme in Industrial Engineering and Management	MASTER'S THESIS

Author: Jonne Viitanen

Subject of the thesis: Estimation of Risk Premium for German Government Bonds Using Regression Based Affine Model

Number of pages: 49	Date: May 26, 2015	Library location: TU
Professorship: Strategic Management		Code of professorship: TU-91

Supervisor: Professor Hannele Wallenius

Instructors: Ruth Kaila, D.Sc. (Tech.) and Lauri Hälikkä, M.Sc.

The purpose of this study is to estimate the risk premium for German government bonds by finding a suitable model to decompose the yield curve. Initially we give an overview on how term structure modeling has evolved starting from expectations hypothesis to the sophisticated affine term structure models.

After discussing several possibilities and weighing strengths and weaknesses we choose to implement a five factor regression based affine term structure model by Adrian, Crump and Moench (2013) that has shown good results with U.S. data. We estimate the model using German yield data extracted from government bonds for the time period between 1997 and 2015.

The model fitted yields and excess returns fit the underlying data excellently. We use the model to forecast future short rates and one month excess returns of long term bonds during the Eurozone crisis (2009 – 2015). Even though the time period is very challenging it performs well at forecasting short rates compared to benchmark models. The model isn't able to anticipate the continuous decline of yields which slightly weakens the results on forecasting excess returns. Finally we show that based on our estimations the short rate expectations have been the main driver of yields except for the past two years when risk premium has been pressed further down to the negative territory.

Keywords: term structure of interest rates; risk premium; forecasting; regression based affine model

Publishing language: English

AALTO-YLIOPISTO	DIPLOMITYÖN TIIVISTELMÄ

Perustieteiden korkeakoulu
Tuotantotalouden koulutusohjelma

Tekijä: Jonne Viitanen		
Työn nimi: Saksan valtionlainojen riskipreemion estimointi regressiopohjaisella affiinilla mallilla		
Sivumäärä: 49	Päiväys: 26.5.2015	Työn sijainti: TU
Professuuri: Strateginen johtaminen		Koodi: TU-91
Työn valvoja: Professori Hannele Wallenius		
Työn ohjaajat: Ruth Kaila, TkT ja Lauri Hälikkä, KTM		

Työn tarkoitus on löytää tehokas malli, jolla estimoida Saksan valtionlainojen riskipreemiota. Aluksi käydään läpi miten korkokäyrän mallintamisessa on kehitytty lähtien odotushypoteesista hienostuneisiin affiineihin malleihin.

Vaihtoehtoisista malleista sopivin on valittu Adrianin, Crumpin ja Moenchin (2013) viiden faktorin regressiopohjainen affiini malli, jolla on saatu hyviä tuloksia U.S. datalla. Malli estimoidaan Saksan datalla aikaperiodilla 1997 – 2015 käyttämällä valtionlainojen nollakuponkikorkoja.

Mallin antamat korot ja ylituotot istuvat erittäin hyvin käytettyyn dataan, eli mallin estimoima korkokäyrä vastaa annettua korkokäyrää. Mallilla ennustetaan tulevia lyhyitä korkoja ja kuukauden ylituottoja pitkän maturiteetin bondeissa eurokriisin aikana (2009 – 2015). Vaikka aikaperiodi on haastava, niin malli ennustaa hyvin tuloksin lyhyitä korkoja. Ylituottojen ennustaminen on hieman haastavampaa, koska malli ei osannut odottaa yhä pieneneviä korkoja pitkissä maturiteeteissa. Silti ennustevirheet ovat vertailukelpoisia.

Lopussa näytetään, että estimoidut odotukset lyhyistä koroista ovat pääasiassa ohjanneet korkoja paitsi aivan viimeisenä kahtena vuotena. Tällöin riskipreemio on painunut nopeasti pakkaselle ja vienyt korot samalla yhä alemmaksi.

Asiasanat: korkokäyrä; riskipreemio; ennustaminen; regressiopohjainen affiini malli	Julkaisukieli: englanti

Acknowledgements

Term structure modeling especially with advanced affine models was very new to me when I started working on this thesis but gradually I was able to master the required knowledge and overcome the obstacles. The journey has been really rewarding but without the support of others I would not have made it this far.

I want to give special thanks to my instructors Lauri Hälikkä and Ruth Kaila who gave me valuable advice and guided me to the right direction. I'm also grateful to my supervisor Professor Hannele Wallenius and to Head of Markets Mika Vihtonen at SEB who made this thesis possible. Tapio Heimo and Patrik Wahlroos deserve my thanks as well for actively helping me to find an interesting and meaningful subject for the thesis.

Finally I want to thank Minna for all the love and support that she's given during all these years.

Jonne Viitanen

Espoo, May 2015

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Background

The risk-free zero-coupon yields calculated from observed bond prices in the market tell us precisely how much we earn annually when holding a bond until maturity. The yields for different maturities can be collected into a yield curve. Previously it was thought that the forward rates extracted from the yield curve implied only the expected future rates by the market (Campbell, 1986). Now we know that the yield curve is influenced also by risk premium (term premium) and convexity bias (Ilmanen, 1995).

The big question is how we can decompose the observed yield curve into these components. Interest rates that are largely driven by expectations of future path of short rates and risk premium form the basis for firms' and consumers' investment and savings decisions. Risk premium, the expected excess return for long term bonds, and convexity help to make better investment decisions and understand the dynamics behind the yield curve. The yield curve contains information about the future path of the economy and term structure modeling can help to find it. (Piazzesi, 2010).

In this study we show how the term structure modeling has evolved over the past decades and discuss several methods to decompose the yield curve into its components. Based on our review we select a model by Adrian, Crump and Moench (2013) and implement it in order to estimate the risk premium (and convexity bias) of German government bonds. Our main focus is the Eurozone crisis that started in late 2009. Its effects are still visible in the European economy (Belkin, et al., 2012; Feldstein, 2015).

1.2. Research Objectives and Questions

The research problem of this thesis can be formulated as following:

How can we effectively estimate the risk premium of German government bonds?

This problem is further divided into the following research questions:

1. *How well the selected method fits the underlying data?*

2. *How well can we forecast bond excess returns during the Eurozone crisis by using the estimated risk premium?*
3. *What is the effect of convexity in determining bond yields?*
4. *According to estimations how much the yields have been driven by changing risk premium and by changing short rate expectations?*

1.3. Structure of the Thesis

The study has five sections of which the first is introduction. The second section explains the basic concepts needed to understand the thesis and a more extensive theory overview on term structure modeling. We discuss yield regressions approach and affine term structure models.

In the third section we explain the reasons why we decided to implement the selected model by Adrian et al. (2013) in order to estimate the risk premium of German government bonds. We show the derivation of the model and how we can extract the model yields so we can build the yield curve for any given point in time. Third section concludes by describing the data we use and what our studied time period is.

The fourth section gives the results as in how we estimated the model parameters, how well the model output fits the underlying input (data) and how well the selected model is able to forecast excess returns (risk premium). Additionally we analyze the dynamics of risk premium over the Eurozone crisis and measure the size of convexity bias. Section five concludes the results of the study and gives future research suggestions based on the findings.

2. Literature Review on Term Structure Modeling

The vast research on term structure modeling started already with the big 20th century economists like Irving Fisher, John Hicks and John Maynard Keynes. Term structure of interest rates describes the cross-section of bond prices at time *t*. Term structure model estimates the term structure of interest rates from a small set of driving factors at any point in time. When we know the dynamics between these factors and the risk premiums, then we can determine the dynamics of the term structure. Term structures of interest rates concentrate on bonds with similar payouts and risks. Most often term structure modeling focuses on risk free zero-coupon bonds. This means bonds are bought at discount to nominal value (if discount rate is positive) and pay the nominal value at maturity without any coupon payments in between. Coupon bond is a portfolio of zero-coupon bonds. Zero-coupon yield curve can be bootstrapped from coupon bonds. (Brandt & Chapman, 2008; Tuckman, 2002). Term structure modeling is important because that way we can derive market expectations of future spot rates and expected term premiums. It is not easy to estimate these future components but it has been shown that current yields contain information about the future path of rates and term premiums. (Fama, 1990).

2.1. Basic Concepts

Bond prices and yields are strictly related to each other as interest rates are derived from traded bond prices. Let us denote by $\ln P_t^{(n)}$ the log price of discount bond with maturity *n* at time *t*. Now the log yield is:

$$y_t^{(n)} = -\frac{1}{n} \ln P_t^{(n)}. \tag{1}$$

By combining the yields for different maturities, we can build a yield curve that shows the cross-section of bond yields at a given point in time *t*. Figure 1 shows an example yield curve.

Figure 1: German zero-coupon yield curve in January 1, 2014.

With risk premium we mean the excess return that investors demand for holding a long term risk-free bond compared to a short term risk-free bond (Cochrane & Piazzesi, 2008). Risk premium is not observable in the markets but it needs to be estimated. The risk premium is caused by the uncertainty in the future path of expected short rates. On average risk premium has been historically positive, meaning that investors require extra compensation for taking risk on future path of short rates. (Ilmanen, 1997; Ilmanen, 1995). We define risk premium as the difference between current yield for n-maturity bond and the expected average short rates (Cochrane & Piazzesi, 2008):

$$rpy_t^{(n)} = y_t^{(n)} - \frac{1}{n} E_t\left(y_t^{(1)} + y_{t+1}^{(1)} \dots + y_{t+n-1}^{(1)}\right), \qquad (2)$$

where E_t is the expectation at time t. In addition to risk premium and future expectations of average short rates, the yield curve is affected by convexity. The convexity effect is usually much smaller than risk premium, and is highly affected by volatility and time to maturity. Simply put convexity refers to the curvature in the yield curve that is caused by the nonlinear relationship between bond price and its yield. As convexity is a valuable property, it tends to press yields lower. If a yield curve is positively convex then a given yield decline rises the price more than a similar rise in yields would lower the price. If we assume constant interest rates for the lifetime of a bond then convexity is the second derivative of bond prices on yield changes:

$$C = \frac{1}{P_t^{(n)}} \frac{d^2 P_t^{(n)}}{{dy_t^{(n)}}^2}. \tag{3}$$

Convexity for bonds is the same as gamma is for options. Convexity bias is the amount of yield that investors give up for getting a positively convex position. The following is an approximation of convexity bias for an *n*-year bond:

$$-\frac{1}{2} * C * \mathrm{Vol}(\Delta y_t^{(n-1)})^2, \tag{4}$$

where $\mathrm{Vol}(\Delta y_t^{(n-1)})^2$ is the squared volatility of yield changes. (Ilmanen, 2000).

2.1.1. Pricing Kernel

A pricing kernel is the stochastic discount factor that is used to discount the future expected payoff to the present. Asset price can be generalized with the following equations:

$$p_t = E(m_{t+1} x_{t+1}), \tag{5}$$

$$m_{t+1} = f(data, parameters). \tag{6}$$

Here p_t is the asset price at time t, x_{t+1} is the payoff at time t+1 and m_{t+1} is the pricing kernel. This definition can be used to price any asset class like for example stocks or bonds. (Cochrane, 2000).

The pricing kernel needs to be defined specifically for the asset in question. The pricing kernel is not known for certainty at time t, therefore it is a stochastic (random) factor. The existence of pricing kernel is implied by the law of one price. Also the existence of strictly positive pricing kernel implies that there are not any arbitrage opportunities. (Cochrane, 2000).

2.1.2. The Expectations Hypothesis

The early dominant term structure model was the expectations hypothesis (Brandt & Chapman, 2008). The pure expectations hypothesis states that risk premium is zero and the more common version of the expectations hypothesis states that risk premium is constant over time (Campbell, 1986). Basically pure expectations hypothesis states that the current forward rates are equivalent to expected future spot rates by the market. Therefore investors do not require a risk premium for holding long-term bonds over short-term bonds. The more common version states that risk premium stays constant over time. Therefore this expectations hypothesis states that investors require always the same risk premium. In this version the forward rates still predict the spot rates when eliminating the effect of constant risk premium. (Campbell, 1986; Bekaert, et al., 1997). There exist few other variations of expectations hypothesis like liquidity preference and preferred habitat theories (Brandt & Chapman, 2008).

Expectations hypothesis has been challenged many times and current view is that it does not properly describe the dynamics of term structure of interest rates (Campbell & Shiller, 1991; Bekaert, et al., 1997; Bekaert, et al., 2007). For example Campbell and Shiller (1991) found that yield spread between long and short rates is correlated with previous short rates: If yield spread is high between long rate and short rate, then long rates usually fall and short rates rise. This is direct contradiction against expectations hypothesis. Bekaert et al. (2007) found that evidence against expectations hypothesis is uniform across different countries and horizons but it can be still good model to analyze effects of monetary policy changes. Conclusion is that time-varying risk premium is a better explanation than expectations theory (Bekaert, et al., 2007; Campbell & Shiller, 1991).

2.2. Yield Regressions

As the evidence is in favor of time-varying risk premium, we first look into a simple regression of estimating risk premium. In this regression approach risk premium is forecasted using forward rates.

Fama and Bliss (1987) already found that spread between *n*-year forward rate and one year yield can predict the risk premium of *n*-year bond. Cochrane and Piazzessi (2005) extended this approach and estimated one year excess returns of *n*-year bond by regressing one year spot rate and forward rates for one year rates up to *n* years. They combined the forward rates into a single factor that allowed to forecast risk premium of bonds up to five years with R^2 as high as 0.44. On average they were able to predict 35 % of variation in future returns. This single factor is also unrelated to the level, slope and curvature factors (Cochrane & Piazzesi, 2005).

Further studies have shown that the predictability of the return forecasting factor found by Cochrane and Piazzessi (CP) (2005) is not that robust across countries and time periods. Extending the time period past bankruptcy of Lehman Brothers and subprime mortgage crisis reduces the predictability R^2 down to 0.25. Also doing the regressions in subsamples show that it is not stable over time, most of the predictability seems to come from the U.S. high inflation period of 1970's and early 1980's. (Wäger, 2012). Still CP return forecasting factor can be used to generate profitable trading strategies as it can forecast excess returns in out-of-sample tests (Kessler & Scherer, 2009).

CP return forecasting factor has shown promising results internationally as well (Kessler & Scherer, 2009; Hellerstein, 2011). In addition to U.S. treasuries the model has shown good results in Australia and Switzerland. Especially good results can be achieved globally if the excess returns are regressed additionally with 1 year U.S. forward rate between 4 and 5 years. Then Germany, Japan, and UK have shown good results as well with only Canada lagging behind. (Kessler & Scherer, 2009). Hellerstein (2011) combined a global forecasting factor from local CP forecasting factors using US, UK, Germany, Japan, Canada, Switzerland, Australia, Sweden, Finland and Norway yield data. Each local forecasting factor was GDP-weighted. This approach allowed to model the flows related to flight-to-quality as in during crisis investors want to invest their money to safe government bonds. For example the Russian default in 1998 resulted in decrease of U.S. risk premiums. (Hellerstein, 2011).

2.3. Affine Term Structure Models

After the rejection of expectations hypothesis, affine term structure models have become the most researched framework for modeling term structures. (Brandt & Chapman, 2008). Affine term structure model means that bond yields are an affine function of some state vector *X*. In other words a linear combination of common factors added with a constant (Piazzesi, 2010):

$$y_t^{(n)} = A_n + B_n' X_t, \tag{7}$$

where A_n is a constant, B_n a vector (K x 1) and X_t a vector of state variables (K x 1). K is the number of pricing factors (state variables). *A* and *B* are solved for each maturity *n*.

The main difference against pure regression based forecasting introduced in last section is the introduction of no-arbitrage that restricts the cross-section dimension (Wäger, 2012). The assumption of no-arbitrage possibilities in affine term structure models implies that there exists a strictly positive pricing kernel (Cochrane, 2000). Affine term structure models are popular because of their tractability as there is no need to calculate yields for example using computationally heavy Monte Carlo simulation. First single factor models had closed-form solutions where the yield was driven completely by short rate factor

(Vasicek, 1977; Cox, et al., 1985). Since then the number of pricing factors has increased and the data-generating process of these factors has evolved. Duffie and Kan (1996) introduced the first framework for multifactor affine term structure models. (Piazzesi, 2010).

Affine term structure models have been studied both in continuous and discrete time settings (Piazzesi, 2010). We concentrate on the discrete time specification of homoscedastic affine term structure models with Gaussian shocks introduced by Ang and Piazzessi (2003). Discrete-time specification does not reduce generality and is often more flexible and tractable than the continuous-time analogue (Joslin, et al., 2011; Le, et al., 2010; Wäger, 2012). The definition of this discrete time affine model class by Ang & Piazzessi (2003) is shown in section 3.2.1 as a part of the derivation of our selected model.

Next we discuss the different affine term structure models that were considered to be used in this study for German data.

2.3.1. Yield Based Affine Term Structure Models

One of the early papers with intention to decompose the yield curve into expected future short rates and risk premiums is Kim and Wright (2005). They fitted a three factor affine model to weekly bond prices and argued that the decline in 10 year Treasury yield around 2004 and 2005 was driven mostly by decline in term premium. Their model was also cited by then sitting Federal Reserve chairman Bernanke who wanted to make clear that risk premium is also driving bond yields and not short rate expectations only. (Kim & Wright, 2005; Cochrane & Piazzesi, 2008)

Cochrane and Piazzessi (2008) introduced a Gaussian affine term structure model based on their return forecasting pricing factor that was introduced in section 2.2. The model has a total of four pricing factors. The first one is the return forecasting pricing factor and the other three are level, slope and curvature that are estimated using the principal component analysis. See Litterman and Scheinkman (1991) for more information about the three principal components. The model has shown good results and can be used to decompose the yield curve into expected average short rates and the risk premium. The drawback of the model is that it has to be estimated using maximum likelihood which is computationally demanding and four factors are not necessarily optimal. (Cochrane & Piazzesi, 2008). There are several extensions and modifications of the model by Cochrane and Piazzessi (2008). Hellerstein (2011) uses the same approach to create an affine model with combining U.S., U.K., Germany and Japan yield data. Similarly combining local and global CP return forecasting factors with level, slope and curvature factors yield promising results (Dahlquist & Hasseltoft, 2013).

It can be criticized that yields are not sufficient by itself to serve as the pricing factors. Much of the variation in bond risk premium cannot necessarily be explained by the cross section of yields but would require other explaining factors, such as macroeconomic factors. (Duffee, 2011).

2.3.2. Utilizing Other Pricing Factors

The models discussed so far have been yields only models. In other words they forecast risk premiums and model the yield curve based on information extracted from yields. Other variations of affine term structure models have been studied as well. One way is to combine survey data into a forecasting factor and add it as an additional pricing factor with the factors extracted from the yields. This might provide more stable results and does not require that long sample for yields. (Kim & Orphanides, 2012). In one version of Cochrane and Piazzessi (2008) realized inflation is regressed against forwards and an inflation forecasting factor is created to estimate the term structure of interest rates (Radwanski, 2012). Similarly unspanned macro factors can be combined to model the term structure (Wäger, 2012; Boos, 2011).

2.3.3. Estimation of Model Parameters

One big problem with the studied affine term structure models is that they use computationally intensive maximum likelihood estimation. This is not very suitable for practitioners if term structure models need to be estimated daily or with short notice for trading purposes. Similarly increasing the number of pricing factors will be difficult. Hamilton and Wu (2012) address this by introducing an affine term structure model that can be estimated with minimum chi-square estimation. They show that it is asymptotically equivalent to maximum likelihood estimation but much easier to compute. With maximum likelihood it is difficult to know if the estimation is a global maximum. The minimum-chi-square estimation allows in some cases to identify straight away if it is a global maximum. (Hamilton & Wu, 2012). Halberstadt and Stapf (2012) use this approach for modeling German bond yields and risk premium.

Adrian, Crump and Moench (2013) introduce a way to estimate affine term structure model with an alternative regression based method. The key is that they do not force bond pricing recursions (see section 3.2.2) in their estimation thus making numerical optimization obsolete. This approach still satisfies the additional restrictions with a very high precision. They use five pricing factors extracted from yields and get very good results with U.S. data when compared against the Cochrane and Piazzessi (2008) four factor specification. (Adrian, et al., 2013).

2.3.4. Shadow Rate Models

Interesting variant of affine term structure models that has been timely topic in the previous year is the shadow rate models. Standard Gaussian affine term structure models do not force the zero lower bound to yields. Shadow rate model proponents argue that zero lower bound or at least only slightly below zero is the absolute floor for rates. This theory has though been challenged in the recent times as interest rates have gone to the negative territory in many developed countries. (Richter & Throckmorton, 2015; Bauer & Rudebusch, 2013; Ichiue & Ueno, 2013).

Basic idea with shadow rate models is that when the model's shadow rates give negative values a zero is used instead. With nonnegative values the shadow rate is used. The further to the negative territory the shadow rate goes, the longer the rate is expected to stay at zero. Ichiue and Ueno (2013) constructed a two factor shadow rate model and compared it against similar affine term structure models. The drawback of their model is the extremely costly estimation as the shadow rate model is nonlinear. (Ichiue & Ueno, 2013). Advances on estimating shadow rate models have been made: Christensen and Rudebusch (2013) were able to estimate a first time three factor shadow rate model by pricing call option on the negative shadow yields.

Zero bound can be forced with other ways as well. For example Quadratic-Gaussian models define yields in quadratic form and define an $n \times n$ symmetric positive-semidefinite matrix that is used to multiply the pricing factors. Probably the best known multifactor affine model that forces the non-negativity of bond yields is the classical CIR model where each factor follows a square root process. (Kim & Singleton, 2012).

3. Selected Model and Data

3.1. Selected Model: Affine Model by Adrian et al. (2013)

As shown on the previous section, there exist numerous promising term structure models with their own strengths and weaknesses. For us it is important that model has proven good results, it is as versatile as possible and relatively fast to estimate.

Top model candidates were the regression model by Cochrane and Piazzessi (2005), four factor affine model by Cochrane and Piazzessi (2008), affine model by Hamilton and Wu (2012), shadow rate model by Bauer and Rudebusch (2013) and regression based affine model by Adrian, Crump and Moench (2013).

Cochrane and Piazzessi's (2005) regression model is the simplest and fastest to estimate of these with relatively good results but does not quite reach the same potential as the others. Four factor affine model by Cochrane and Piazzessi (2008) nicely estimates the yield curve and risk premia but it uses maximum likelihood for estimation which is computationally demanding and not that suited for practical uses. Hamilton and Wu (2012) introduced similar affine term structure model as the one Cohrane and Piazzessi (2008) used but they estimate the model with faster minimum chi-square estimation which still requires demanding numerical optimization. Shadow rate model by Bauer and Rudebusch (2013) is theoretically interesting as it forces the zero lower bound for yields. However recent monetary actions have proven this limit to be questionable. We selected the five factor model by Adrian, Crump and Moench (2013) as it combines extremely fast estimation of regression models into sophisticated framework of affine models with very good results in the U.S. data.

One advantage of the selected model is that we can also use observable pricing factors like macroeconomic variables to estimate the model. This is contrarian to many affine models that require the pricing factors to be extracted only from yields. Other big advantage of this model is that we do not strictly need zero coupon yield curve to estimate the model, we can use other fixed income securities as well without bootstrapping zero coupon yield curve (Adrian, et al., 2013). This is especially useful for practical purposes when we want to use other tradable instruments. Also our study contributes new research as the model has not been previously applied for German data.

3.2. The Model

This section shows the derivation of the model for analyzing the risk premium of German government bonds. The model is the one introduced by Adrian et al. (2013). Model uses an exponentially affine pricing kernel as discussed earlier. The parameters are estimated using three linear regressions. (Adrian, et al., 2013).

3.2.1. Derivation of Excess Returns

We follow the approach of Adrian et al. (2013) for deriving the model. The following vector autoregression describes the evolution of state variables X_t (K x 1 vector):

$$X_{t+1} = \mu + \Phi \mathrm{X_t} + v_{t+1}, \tag{8}$$

where μ is mean (K x 1 vector), Φ is a matrix (K x K) and v_{t+1} are the shocks (K x 1 vector). Vector autoregressive models can be estimated using ordinary least squares (OLS). Forecasting with these models is straightforward by continuing recursively from the end of the period onwards. (Verbeek, 2004). The shocks v_{t+1} are assumed to follow a Gaussian distribution with covariance matrix Σ:

$$v_{t+1} | \{X_s\}_{s=0}^{t} \sim N(0, \Sigma), \tag{9}$$

where $\{X_s\}_{s=0}^{t}$denotes the history of $\mathrm{X_t}$. The no-arbitrage assumption implies that there exists a following pricing kernel M_t:

$$P_t^{(n)} = E_t\left[M_{t+1}P_{t+1}^{(n-1)}\right]. \quad (10)$$

In other words, the price of n-year bond is the expected value of the same bond one year from now discounted to present with the stochastic discount factor M_{t+1}. The pricing kernel M_{t+1} is assumed to be exponentially affine:

$$M_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\Sigma^{-1/2}v_{t+1}\right), \quad (11)$$

where $r_t = \ln P_t^{(1)}$ is the continuously compounded risk-free rate and λ_t is the market prices of risk. The math used to derive the affine pricing kernel derivation is shown by Ang and Piazzessi (2003). The market prices of risk are assumed to be essentially affine as defined by Duffee (2002).

$$\lambda_t = \Sigma^{-\frac{1}{2}}(\lambda_0 + \lambda_1 X_t). \quad (12)$$

The log excess holding period return for a bond maturing in n periods:

$$rx_{t+1}^{(n-1)} = \ln P_{t+1}^{(n-1)} - \ln P_t^{(n)} - r_t. \quad (13)$$

Using Equation (11) and Equation (13) in Equation (10) gives

$$1 = E_t\left[\exp\left(rx_{t+1}^{(n-1)} - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\Sigma^{-\frac{1}{2}}v_{t+1}\right)\right]. \quad (14)$$

$\left\{rx_{t+1}^{(n-1)}, v_{t+1}\right\}$ are assumed to be jointly normally distributed. Now we get

$$E_t\left[rx_{t+1}^{(n-1)}\right] = Cov_t\left[rx_{t+1}^{(n-1)}, v'_{t+1}\Sigma^{-\frac{1}{2}}v_{t+1}\right] - \frac{1}{2}Var_t\left[rx_{t+1}^{(n-1)}\right]. \tag{15}$$

We denote

$$\beta_t^{(n-1)\prime} = Cov_t\left[rx_{t+1}^{(n-1)}, v'_{t+1}\right]\Sigma^{-1}, \tag{16}$$

And using Equation (12) we get

$$E_t\left[rx_{t+1}^{(n-1)}\right] = \beta_t^{(n-1)'}(\lambda_0 + \lambda_1 X_t) - \frac{1}{2}Var_t\left[rx_{t+1}^{(n-1)}\right]. \tag{17}$$

Now we decompose the unexpected excess return into two components. One is correlated with v_{t+1} and the other is conditionally orthogonal. Now we find

$$rx_{t+1}^{(n-1)} - E_t\left[rx_{t+1}^{(n-1)}\right] = \gamma_t^{(n-1)'}v_{t+1} + e_{t+1}^{(n-1)}, \tag{18}$$

where $e_{t+1}^{(n-1)}$ is the return pricing error and $\gamma_t^{(n-1)'}$ is a vector (1 x K). Using Equation (16) shows that $\gamma_t^{(n-1)} = \beta_t^{(n-1)}$. It is assumed that return pricing errors

$e_{t+1}^{(n-1)}$ are conditionally independently and identically distributed (i.i.d.) and have a variance σ^2.

Linear combinations of log yields are used as observable factors $\mathrm{X_t}$. The model parameters are estimated using holding period returns based on the same set of yields. This implies that $\beta_t = \beta \forall t. \beta$ is assumed to be constant.

Now we get the return generating process for log excess holding period returns:

$$rx_{t+1}^{(n-1)} = \beta^{(n-1)\prime}(\lambda_0 + \lambda_1 X_t) - \frac{1}{2}\left(\beta^{(n-1)\prime} \Sigma \beta^{(n-1)} + \sigma^2\right) + \beta^{(n-1)\prime} v_{t+1} + e_{t+1}^{(n-1)}. \tag{19}$$

The first term is expected return, second is convexity adjustment, third is priced return innovation and last term is return pricing error. Return innovation is the explainable forecast error of expected returns. Transforming this across maturities and time periods, we can write it as

$$rx = \beta'(\lambda_0 \iota_T' + \lambda_1 X_-) - \frac{1}{2}(B^* vec(\Sigma) + \sigma^2 \iota_N)\iota_T' + \beta' V + E, \tag{20}$$

where rx is a matrix (N x T) of excess returns, $\beta = [\beta^{(1)} \beta^{(2)} \ldots \beta^{(N)}]$ is a matrix (K x N) of factor loadings, ι_N and ι_T are vectors (N x 1 and T x 1) of ones, $X_- = [X_0 \, X_1 \ldots X_{T-1}]$ is a matrix (K x T) of lagged pricing factors, $B^* = [vec(\beta^{(1)}\beta^{(1)\prime}) \ldots vec(\beta^{(N)}\beta^{(N)\prime})$ is an N x K^2 matrix, V is a K x T matrix and E is an N x T matrix.

Vec is the vectorization of a matrix. In other words it is a linear transformation of a matrix into a column vector where each column of a matrix is stacked above each other.

3.2.2. Model Yields

When the model parameters are estimated, it allow us to generate a zero coupon yield curve. We follow the approach of Adrian et al. (2013) for extracting the affine yields.

It can be shown that bond prices are exponentially affine in the vector of state variables based on the previous assumptions:

$$\ln P_t^{(n)} = A_n + B_n' X_t + u_t^{(n)}, \tag{21}$$

where $u_t^{(n)}$ is the log yield pricing error. Equation (21) is substituted into Equation (13), when we get:

$$\begin{aligned} rx_{t+1}^{(n-1)} = A_{n-1} + B_{n-1}' X_{t+1} + u_{t+1}^{(n-1)} - A_n - B_n' X_t - u_t^{(n)} \\ + A_1 + B_1' X_t + u_t^{(1)}. \end{aligned} \tag{22}$$

Now we can substitute the above with the return generating expression in Equation (19), when we get:

$$\begin{aligned} A_{n-1} + B_{n-1}'(\mu + \Phi \mathrm{X_t} + v_{t+1}) + u_{t+1}^{(n-1)} - A_n - B_n' X_t \\ - u_t^{(n)} + A_1 + B_1' X_t + u_t^{(1)} \\ = \beta^{(n-1)'}(\lambda_0 + \lambda_1 X_t + v_{t+1}) \\ - \frac{1}{2}\left(\beta^{(n-1)'} \Sigma \beta^{(n-1)} + \sigma^2\right) + e_{t+1}^{(n-1)}. \end{aligned} \tag{23}$$

This must hold state by state. Let $A_1 = -\delta_0$ and $B_1 = -\delta_1$ that are estimated by regressing the one-month yield on pricing factors:

$$\Delta = r_t[\iota_T X]'([\iota_T X][\iota_T X]')^{-1}, \tag{24}$$

when $\delta_0 = \Delta_1$ and $\delta_1 = [\Delta_2 \dots \Delta_K]$. Now we can get the following recursive linear equations for the bond pricing parameters:

$$A_n = A_{n-1} + B'_{n-1}(\mu - \lambda_0) + \frac{1}{2}(B'_{n-1}\Sigma B_{n-1} + \sigma^2) - \delta_0, \tag{25}$$

$$B'_n = B'_{n-1}(\Phi - \lambda_1) - \delta'_1, \tag{26}$$

$$A_0 = 0, B'_0 = 0, \tag{27}$$

$$\beta^{(n)'} = B'_n. \tag{28}$$

Log bond pricing errors can be expressed now with the following equation:

$$u_{t+1}^{(n-1)} - u_t^{(n)} + u_t^{(1)} = e_{t+1}^{(n-1)}, \tag{29}$$

where left side of the equation is the log yield pricing error and the right side is the return pricing error. There are a few noteworthy things to notice. The log bond prices derivation is exact as long as $\beta^{(n)'} = B'_n$. Equation (25) and Equation

(26) are almost the same as standard linear difference equations with homoscedastic shocks for affine term structure models. The only difference is that there is an additional $\frac{1}{2}\sigma^2$ term in Equation (25). The term is included because this approach includes the pricing errors in the no-arbitrage recursions directly. (Adrian, et al., 2013).

If log yield pricing errors are i.i.d. then return pricing errors are serially and cross-sectionally correlated as implied by Equation (29). Traditional maximum likelihood based approach for affine term structure models assume serially uncorrelated yield pricing errors that implies return pricing errors to be serially correlated. (Adrian, et al., 2013).

We can easily generate the risk neutral bond pricing parameters A_n^{RF} and B_n^{RF} by setting the risk parameters λ_0 and λ_1 in Equation (25) and Equation (26) to zero. This way we can generate risk-neutral yields directly by using the formula:

$$-\frac{1}{n}\left(A_n^{RF} + B_n^{RF'} X_t\right), \tag{30}$$

where A_n^{RF} is a constant and B_n^{RF} is a vector (K x 1). This equation is the time t expected average future short rates for the next n periods. The risk premium can be calculated by deducting the risk neutral yield from the model implied fitted yield. This risk premium is the expected excess return for longer term bonds compared to short term bonds. It is the expected return when buying n year bonds that are funded by selling short term bonds. (Adrian, et al., 2013). This model implied risk premium will be used for studying the risk premium of German government bonds during the recent years.

3.3. Data and Time Period

We use zero coupon yield data constructed by Bundesbank from listed Federal securities. Bundesbank estimates the term structure using Nelson-Siegel-Svensson method (Schich, 1997). Nelson and Siegel (1987) fit a discount function for the term structure that can be used to solve any forward or spot rate for any given maturity. This way the whole yield curve can be constructed by using a few specified parameters (Nelson & Siegel, 1987). The one year forward rate in *n* years according to Nelson and Siegel (1987):

$$f_t^{(n)} = \beta_0 + \beta_1 e^{-\frac{n}{\tau_1}} + \beta_2 \frac{n}{\tau_1} e^{-\frac{n}{\tau_1}}, \tag{31}$$

where $\beta_0,\ \beta_1,\ \beta_2$ and τ_1 are constants at time *t*. Svensson (1994) extends this by adding a fourth term that improves fit and increases flexibility. Then the one year forward rate in *n* years is the following (Svensson, 1994):

$$f_t^{(n)} = \beta_0 + \beta_1 e^{-\frac{n}{\tau_1}} + \beta_2 \frac{n}{\tau_1} e^{-\frac{n}{\tau_1}} + \beta_3 \frac{n}{\tau_2} e^{-\frac{n}{\tau_2}}, \tag{32}$$

where τ_2 is constant at time *t*. Now we can calculate the zero rates by integrating the forward rates (Svensson, 1994):

$$\begin{aligned} y_t^{(n)} = \beta_0 + \beta_1 \left(\frac{1 - e^{-\frac{n}{\tau_1}}}{\frac{n}{\tau_1}} \right) + \beta_2 \left(\frac{1 - e^{-\frac{n}{\tau_1}}}{\frac{n}{\tau_1}} - e^{-\frac{n}{\tau_1}} \right) \\ + \beta_3 \left(\frac{1 - e^{-\frac{n}{\tau_2}}}{\frac{n}{\tau_2}} - e^{-\frac{n}{\tau_2}} \right). \end{aligned} \tag{33}$$

In the functions n is the maturity, $\beta_0, \beta_1, \beta_2, \beta_3, \tau_1$ and τ_2 are the parameters to be solved. There are two common ways to find solutions for these parameters. One is to minimize the sum of squared deviations between the estimated and observed bond prices. However this approach may result in substantial yield errors at the short end of the yield curve. The other way that Bundesbank also uses is to minimize yield errors instead. This is computationally more demanding but fits well with the objective of studying yields. (Svensson, 1994; Schich, 1997). The mean yield error using Nelson-Siegel-Svensson method was 11.99 basis points according to Schich (1997) for period 1972 – 1996 when estimating the German term structure.

Bundesbank uses three types of listed Federal securities for building the zero yield curve. Keeping the number of different types small is important to make sure that the included securities are as homogeneous as possible. Still a large number of securities are needed so that all maturity segments can be estimated as reliably as possible. (Schich, 1997). With these in mind Bundesbank uses Federal bonds (Bunds), five-year Federal notes (Bobls) and Federal Treasury notes (Schätze) (Bundesbank, 2015; Schich, 1997).

The yield data from Bundesbank is from 1997 onwards. The model is estimated from August 29, 1997 to March 31, 2015 for the model fit tests, convexity bias estimation and qualitative analysis of risk premium dynamics. The main focus of qualitative analysis is still on the Eurozone crisis period.

During the out-of-sample forecasting we concentrate on the Eurozone crisis, therefore the model is initially estimated for the sample between August 29, 1997 and August 31, 2009. Then the model is re-estimated on monthly basis by expanding the estimation sample up until March 31, 2015.

4. Results

In this section we estimate the German yield curve and decompose it into risk premium and expected short rates. We also analyze how the risk premium of German government bonds behaved during the Eurozone crisis and especially how well the selected model was able to predict the excess returns.

4.1. Model Specification

Adrian et al. (2013) find that using five principal components of yields outperforms three and four component specifications when analyzing U.S. Treasury yield curves. The comparison is done using the Anderson (1951) canonical correlations test as well with the Wald statistic with the hypothesis that whether the last column of β' is zero. Both tests support that the five factor specification gives likely a better fit for bond returns and risk premiums compared to using fewer factors. (Adrian, et al., 2013). Therefore we opt to use a five factor specification as well.

As the model is regression based, it does not make the five factor specification computationally any more demanding than specifications with fewer factors (Adrian, et al., 2013). Many term structure models impose the nonlinear cross-section restrictions that make the estimation much more demanding when increasing the number of factors (Piazzesi, 2010).

4.2. Estimation of the Model

We follow the approach of Adrian et al. (2013) for estimating the model. The following three step regression is used to estimate the parameters of the model.

1. Initially we solve the pricing factors *X* from observed yields using principal component analysis. It is a powerful method to decompose a set of observations into uncorrelated (orthogonal) components (Jolliffe, 2002). Next we estimate Equation (8) using Ordinary Least Squares (Verbeek, 2004):

$$\widehat{\Phi} = XX'_{-}(X_{-}X'_{-})^{-1}, \tag{34}$$

where $X_{-} = [X_0\ X_1 \ldots X_{T-1}]$ and $X = [X_1\ X_2 \ldots X_T]$. This way the X_{t+1} can be decomposed into a predictable component as well as into an estimate of $\hat{v}_{t+1}$ (pricing factor innovation). These are stacked into the matrix $\widehat{V}$:

$$\widehat{\mathrm{V}} = X - \widehat{\Phi}X_{-}. \tag{35}$$

$\widehat{V}$ will be used to construct an estimator of covariance matrix $\widehat{\Sigma} = \widehat{V}\widehat{V}'/T$.

2. Excess returns are regressed on a constant, lagged pricing factors and contemporaneous pricing factor innovations:

$$rx = a\iota'_T + \beta'\widehat{V} + cX_{-} + E. \tag{36}$$

Regressors are collected into the (2K + 1) x T matrix $\tilde{Z} = \left[\iota_T\ \widehat{V}'\ X'_{-}\right]'$:

$$rx = [\widehat{\boldsymbol{a}}\ \beta'\ \widehat{\boldsymbol{c}}]\tilde{Z} + E. \tag{37}$$

Now the estimators become

$$[\widehat{\boldsymbol{a}}\ \beta'\ \widehat{\boldsymbol{c}}] = rx\tilde{Z}'\left(\tilde{Z}\,\tilde{Z}'\right)^{-1}. \tag{38}$$

Residuals from this regression are collected into a matrix $\hat{E}$ (N x T). Now $\hat{\sigma}^2 = \frac{\mathrm{trace}(\hat{E}\hat{E}')}{NT}$. $\hat{B}^*$ is constructed from $\hat{\beta}$:

$$B^* = [vec(\beta^{(1)}\beta^{(1)\prime}) \dots vec(\beta^{(N)}\beta^{(N)\prime}). \tag{39}$$

3. The price of risk parameters λ_0 and λ_1 are estimated using cross-sectional regression. Equation (20) tells that $\boldsymbol{a} = \beta'\lambda_0 - \frac{1}{2}(B^*vec(\Sigma) + \sigma^2\iota_N)$ and $\boldsymbol{c} = \beta'\lambda_1$. These expressions are used to obtain the following estimators:

$$\hat{\lambda}_0 = (\hat{\beta}\hat{\beta}')^{-1}\hat{\beta}\left(\hat{\boldsymbol{a}} + \frac{1}{2}(B^*vec(\Sigma) + \sigma^2\iota_N)\right), \tag{40}$$

$$\hat{\lambda}_1 = (\hat{\beta}\hat{\beta}')^{-1}\hat{\beta}\hat{\boldsymbol{c}}. \tag{41}$$

4.3. Model Fit

Adrian et al. (2013) found in their study that the model provided extremely good fit for the United States data between 1987 and 2011 for maturities up to 10 years. We examine how well the model fits the German yield curve between 1997 and 2015.

First we examine the yield pricing errors as in how well the model output fits the inputted zero coupon yield curve. This zero coupon yield curve is constructed using Nelson-Siegel-Svenson model with the parameters published by Bundesbank. Model fitted yields are constructed via the recursive bond pricing parameters that are solved using Equation (25), Equation (26) and Equation (27). This way we can calculate fitted yields for all maturities at monthly intervals and can calculate yield pricing errors:

$$\hat{u}_t^{(n)} = \ln P_t^{(n)} + \frac{1}{n}\left(\hat{A}_n + \hat{B}_n' X_t\right). \quad (42)$$

We estimate the yield pricing errors for maturities up to ten years. Period is from August 29, 1997 to March 31, 2015. Table 1 and Table 2 show the calculated statistics: mean, standard deviation, skewness and kurtosis of the sample.

Table 1: Statistics for yield pricing errors for maturities from 1 year to 5 years, n = months

Statistics	n = 12	n = 24	n = 36	n = 48	n = 60
Mean	-0.011	-0.007	-0.006	-0.006	-0.005
Standard deviation	0.024	0.012	0.018	0.020	0.018
Skewness	0.044	-0.177	-0.458	-0.629	-0.640
Kurtosis	2.408	0.519	2.048	1.500	0.156

Table 2: Statistics for yield pricing errors for maturities from 6 years to 10 years, n = months

Statistics	n = 72	n = 84	n = 96	n = 108	n = 120
Mean	-0.002	0.001	0.003	0.003	0.002
Standard deviation	0.015	0.012	0.009	0.009	0.016
Skewness	-0.833	-0.965	-0.966	-0.869	-0.731
Kurtosis	-0.215	-0.137	0.014	1.764	1.812

As the above statistics show the average yield pricing error for all maturities is around one basis point or less. Standard deviations show that the variation in errors is also very small. Yield pricing errors are on average skewed more to the negative values thus meaning that on average fitted yields are slightly overestimated. Kurtosis is for most maturities positive thus indicating a higher peak and fatter tails for the errors. However kurtosis differs significantly between maturities.

Figure 2 and Figure 3 show the fit for 2 year and 10 year maturities for the time period. Additionally we chart the term premium that is the difference between fitted yields and risk neutral yields that we derived in section 3.2.2:

$$-\frac{1}{n}(A_n + B_n' X_t) + \frac{1}{n}\left(A_n^{RF} + B_n^{RF'} X_t\right). \tag{43}$$

The fit for yields is extremely good considering that the model is fitted for returns.

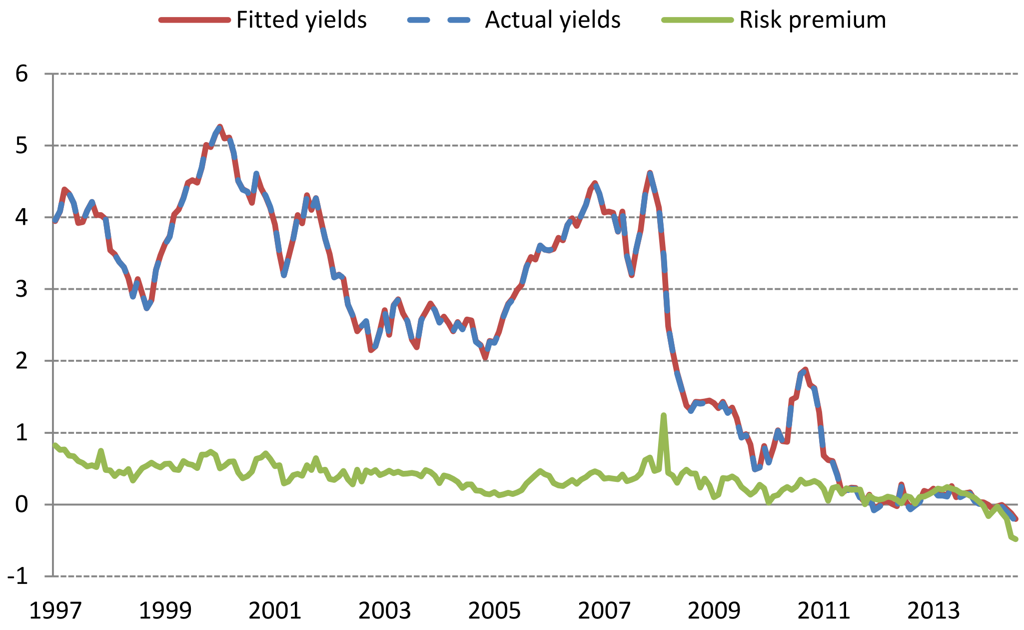

Figure 2: Actual vs. model fitted yields and term premium for n = 24 months

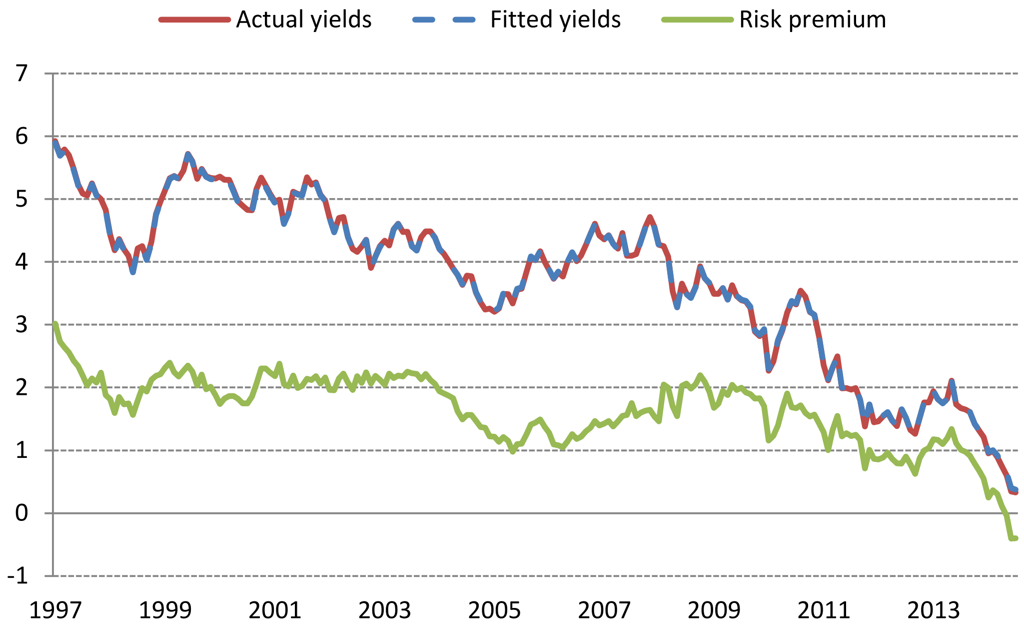

Figure 3: Actual vs. model fitted yields and term premium for n = 120 months

Second we examine how well the model given fitted excess returns match actual excess returns. In other words how large the return pricing errors are. We use the same period from August 29, 1997 to March 31, 2015. We get fitted excess returns from Equation (19) so now we can calculate return pricing errors as the difference of actual excess returns and fitted excess returns:

$$\hat{e}_{t+1}^{(n-1)} = \ln P_{t+1}^{(n-1)} - \ln P_t^{(n)} - r_t - (\hat{B}_{n-1}'(\hat{\lambda}_0 + \hat{\lambda}_1 X_t) - \frac{1}{2}(\hat{B}_{n-1}'\Sigma\hat{B}_{n-1} + \hat{\sigma}^2) + \hat{B}_{n-1}'\hat{v}_{t+1}). \tag{44}$$

Table 3 and Table 4 show the calculated statistics: mean, standard deviation, skewness and kurtosis of the sample.

Table 3: Statistics for return pricing errors for maturities from 1 year to 5 years, n = months

Statistics	n = 12	n = 24	n = 36	n = 48	n = 60
Mean	-3.39E-05	-2.83E-05	-2.31E-05	-1.06E-05	5.23E-06
Standard deviation	0.002	0.005	0.008	0.011	0.013
Skewness	0.200	0.053	0.048	0.096	0.142
Kurtosis	1.435	0.544	0.468	0.268	0.131

Table 4: Statistics for return pricing errors for maturities from 6 year to 10 years, n = months

Statistics	n = 72	n = 84	n = 96	n = 108	n = 120
Mean	1.92E-05	3.13E-05	4.44E-05	6.21E-05	8.78E-05
Standard deviation	0.016	0.018	0.021	0.023	0.026
Skewness	0.163	0.157	0.132	0.096	0.054
Kurtosis	0.095	0.142	0.242	0.365	0.496

Return pricing errors are almost non-existent for all maturities. Variance is very low and return pricing errors are slightly skewed to positive values. Kurtosis is close to zero. In Figure 4 and Figure 5 shows the whole time-series fit for two year and ten year maturities. In addition we show the expected excess returns that we get as the sum of expected return component and convexity adjustment component but leaving out the priced return innovation (explainable forecast error):

$$\hat{B}'_{n-1}\left(\hat{\lambda}_0 + \hat{\lambda}_1 X_t\right) - \frac{1}{2}\left(\hat{B}'_{n-1}\Sigma\hat{B}_{n-1} + \hat{\sigma}^2\right). \tag{45}$$

We can conclude that the model excess returns give excellent fit against actual excess returns.

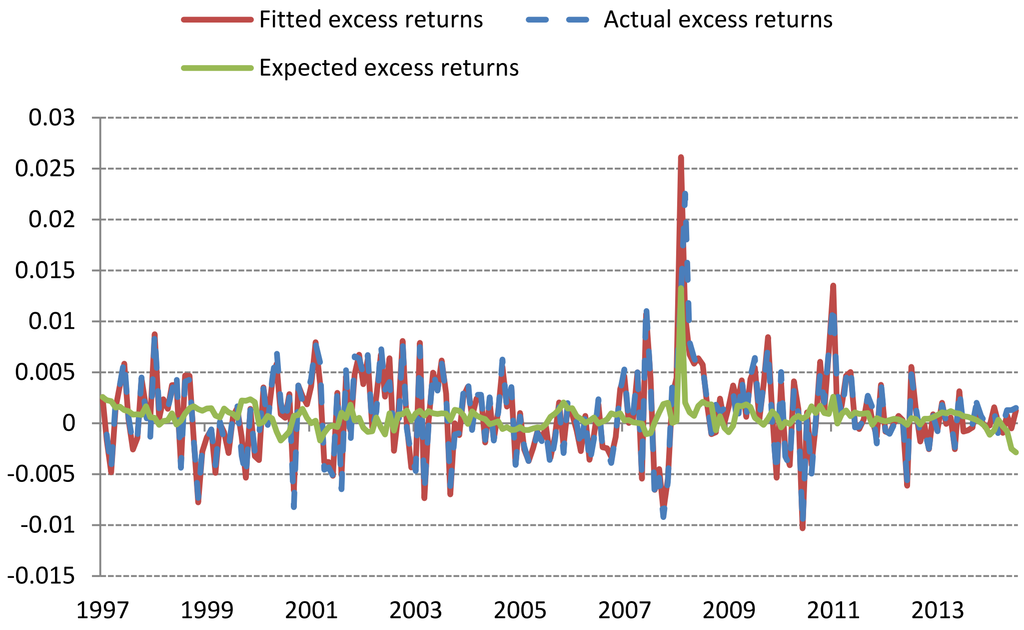

Figure 4: Actual vs. model fitted excess returns and expected returns for n = 24 months

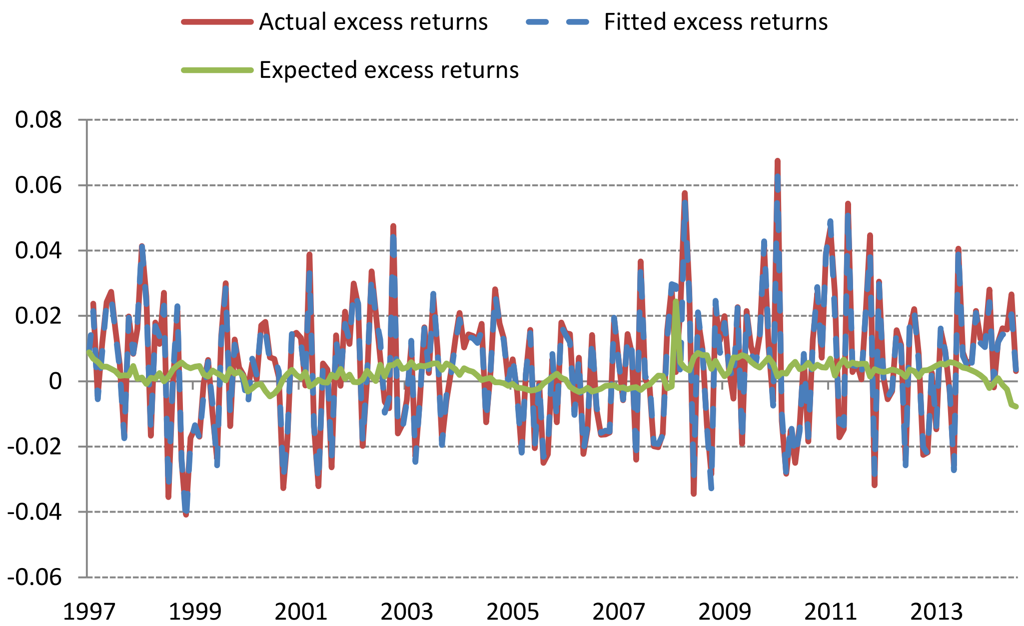

Figure 5: Actual vs. model fitted excess returns and expected returns for n = 120 months

All in all, we find similar results as Adrian et al. (2013) found for U.S. Treasuries. Model fit is very good even though the cross-equation recursions in Equation

(25) and Equation (26) are not imposed during estimation but used only for calculating yields for all the maturities.

4.4. Out-of-Sample Forecasting

In this section we investigate how well the model is able to forecast the future path of short term interest rates and bond excess returns. This way we can compare how reliable the risk premium estimate is for different maturities. First we study how well the model is able to forecast short term rates up to three years. Second we test how accurately the model can forecast one month excess returns for maturities up to ten years.

4.4.1. Expected Future Short Rates

Equation (2) defines the risk premium as the difference between current yield for n-maturity bond and average expected short rates. This means that the difference between realized average short rates and expected short rates is equal to the difference between expected excess return (risk premium) and realized excess return because the yield at time t is always known. We ignore convexity bias as it is not significant at the short end of the yield curve (see section 4.5).

We use the period between September 1, 1997 and August 31, 2009 for German government bonds as the basis for the out-of-sample forecasting. Model is initially estimated for that period and after that re-estimated on monthly basis up until March 31, 2015. We look at the short rate expectations given by the model for 6 month, 1 year, 2 year and 3 year maturities each month and compare it to the realized average one month short rates for the lifetime of the bond. We use Equation (30) to calculate the average expected short rates. The time period between August 31, 2009 and March 31, 2015 is when the European debt crisis unfolded and is thus an interesting and at the same time very challenging out-of-sample testing period for this study.

The studied model by Adrian et al. (2013) is compared against two simple benchmark models. The first benchmark model is a risk premium hypothesis (basically a random walk) where we estimate that the best guess for future interest rate is the rate today. Therefore the expected average one month interest rate for up to three years is estimated to be the current one month rate. In risk premium hypothesis the yield curve is expected to remain unchanged and forwards reflect only expected excess returns. This is the opposite definition to pure expectations theory where forwards reflect the changing interest rate expectations. Truth is somewhere between these but risk premium hypothesis has been proven to be empirically more correct than pure expectations hypothesis. (Ilmanen, 1996). The other benchmark model is based on historical short rates. We use the average realized one month short rates for the past three years as the expected average one month short rate for up to three years.

We calculate the root-mean-square deviation (RMSD) for all models during the out-of-sample period for maturities up to three years:

$$RMSD = \sqrt{\frac{\sum_{t=1}^{m}\left(\frac{1}{n}E_t\left(y_t^{(1)}+y_{t+1}^{(1)}\dots+y_{t+n-1}^{(1)}\right)-\frac{1}{n}\left(y_t^{(1)}+y_{t+1}^{(1)}\dots+y_{t+n-1}^{(1)}\right)\right)^2}{m}}, \tag{46}$$

where m is the number of observations. The results for forecasting average short rates for 6 months, 1 year, 2 years and 3 years are shown in Table 5, Table 6 and Table 7.

Table 5: Out-of-sample forecasting RMSD for forecasting future average short rates with model by Adrian et al. (2013)

Maturity (months)	Observations (months)	RMSD
6	62	0.261 %
12	56	0.424 %
24	44	0.618 %
36	32	0.858 %

Table 6: Out-of-sample forecasting RMSD for forecasting future average short rates with benchmark model (risk premium hypothesis)

Maturity (months)	Observations (months)	RMSD
6	62	0.315 %
12	56	0.385 %
24	44	0.466 %
36	32	0.576 %

Table 7: Out-of-sample forecasting RMSD for forecasting future average short rates with benchmark model (average historical rates)

Maturity (months)	Observations (months)	RMSD
6	62	0.865 %
12	56	0.910 %
24	44	1.021 %
36	32	1.175 %

The model by Adrian et al. (2013) clearly outperforms the second benchmark model that uses average historical rates. However the first benchmark model (risk premium hypothesis) is better in all maturities up to three years except for forecasting short rates up to 6 months. As expected accuracy decreases for all models when maturity increases, as future path of short rates becomes more uncertain.

The results for the German data between August 31, 2009 and March 31, 2015 are significantly better than the results obtained by Adrian et al. (2013) for U.S. data between January, 1992 and December, 2011. Their average RMSD ranged from bit less than 0.4 % for 6 months to roughly 1.5 % for 3 years. However the model was better than random walk in all maturities up to 60 months that they studied. (Adrian, et al., 2013). With German data the model has difficulties against expectations of unchanging short rates. Our sample size is lot smaller and characterized by highly unusual period of low rates compared to history because of the Eurozone crisis. Short rates have stayed low and relatively stable thus making the expectation of constant short rates (random walk) a very good forecast that reflects the very good results in Table 6. Also the time period between 2009 and 2015 is very different from 1997 – 2009 for German data which is analyzed in section 4.5. This gives the affine model a disadvantage because it uses all the information observed in yields from 1997 onwards.

4.4.2. Expected One Month Excess Returns

The longer maturity bonds are not taken into account in the previous out-of-sample forecasting. Therefore we look into how well the model by Adrian et al. (2013) is able to forecast one month excess returns for bond maturities up to 10 years. This is especially interesting for investors because it is important to know how reliable the risk premium estimates for different maturities are.

We get the model estimated expected excess return as a sum of expected return component and convexity adjustment component from Equation (19):

$$E_t\left[rx_{t+1}^{(n-1)}\right] = \beta^{(n-1)\prime}(\lambda_0 + \lambda_1 X_t) - \frac{1}{2}\left(\beta^{(n-1)\prime}\Sigma\beta^{(n-1)} + \sigma^2\right). \quad (47)$$

The model by Adrian et al. (2013) is again initially estimated for period between September 1, 1997 and August 31, 2009 using German data. We compare the model against two benchmark models data between August 31, 2009 and March 31, 2015 on monthly basis. The first benchmark model is the risk premium hypothesis introduced in the previous section. In other words expected future interest rate for any maturity is the same as today's rate for that maturity. Now we get the expected one month excess return for this benchmark model as:

$$E_t\left[rx_{t+1}^{(n-1)}\right] = \ln P_t^{(n-1)} - \ln P_t^{(n)} - r_t. \quad (48)$$

The second benchmark model estimates the one month expected excess return for each maturity as the average realized one month excess return for that maturity for the past five years. Now we calculate the RMSD of expected one month excess returns against the realized one month excess returns for maturities up to 10 years for each model:

$$RMSD = \sqrt{\frac{\sum_{t=1}^{m}\left(E_t\left(rx_{t+1}^{(n-1)}\right) - rx_{t+1}^{(n-1)}\right)^2}{m}}, \tag{49}$$

Where m = 67 (monthly observations). Results are shown in Table 8 and Table 9. The risk premium hypothesis benchmark model has the best overall results. Historical benchmark model is worse than the model by Adrian et al. (2013) in short maturities but gets better in long maturities. However the models are quite close to each other with no significant differences based on root-mean-square deviations.

Table 8: Out-of-sample RMSD for forecasting one month excess returns for maturities between 1 and 5 years for the period between August 31, 2009 and March 31, 2015 (n = months)

Model	n = 12	n = 24	n = 36	n = 48	n = 60
Model by Adrian et al. (2013)	0.132 %	0.330 %	0.556 %	0.793 %	1.036 %
Risk premia hypothesis	0.125 %	0.322 %	0.544 %	0.768 %	0.990 %
Historical excess returns	0.140 %	0.341 %	0.564 %	0.787 %	1.005 %

Table 9: Out-of-sample RMSD for forecasting one month excess returns for maturities between 6 and 10 years for the period between August 31, 2009 and March 31, 2015 (n = months)

Model	n = 72	n = 84	n = 96	n = 108	n = 120
Model by Adrian et al. (2013)	1.282 %	1.532 %	1.788 %	2.049 %	2.316 %
Risk premia hypothesis	1.210 %	1.432 %	1.657 %	1.887 %	2.122 %
Historical excess returns	1.222 %	1.439 %	1.658 %	1.882 %	2.112 %

The realized one month excess returns have been exceptionally large during 2009 – 2015 for long term bonds. Table 10 shows the realized one month excess returns converted into realized annualized risk premiums over the one month interest rate. The longer the maturity, the higher the excess returns have been as yields over the curve have been pressed down. For 10 year bonds the realized one month average excess return (risk premium) has been more than double between September 2009 and March 2015 compared to period between September 1997 and August 2009.

Table 10: Average realized one month annualized risk premiums (%) for 1 year, 3 year, 5 year, 7 year and 10 year bonds between 1997 and 2015.

Time period	n = 12	n = 36	n = 60	n = 84	n = 120
Sep 1997 - Aug 2009	0.57 %	1.67 %	2.59 %	3.35 %	4.24 %
Sep 2009 - Mar 2015	0.36 %	2.17 %	4.40 %	6.51 %	9.27 %

The exceptional period of Eurozone crisis has been difficult for the studied model by Adrian et al. (2013) when using the historical yields since 1997 as an input because the yields have declined and the German yield curve has flattened and stayed flat. This is very different compared to historical levels and shapes of German yield curve. See Figure 6 for comparison. The model has been expecting on average slightly negative one month excess returns for long term bonds but instead the yields have been going further down and earning highly positive one month excess returns.

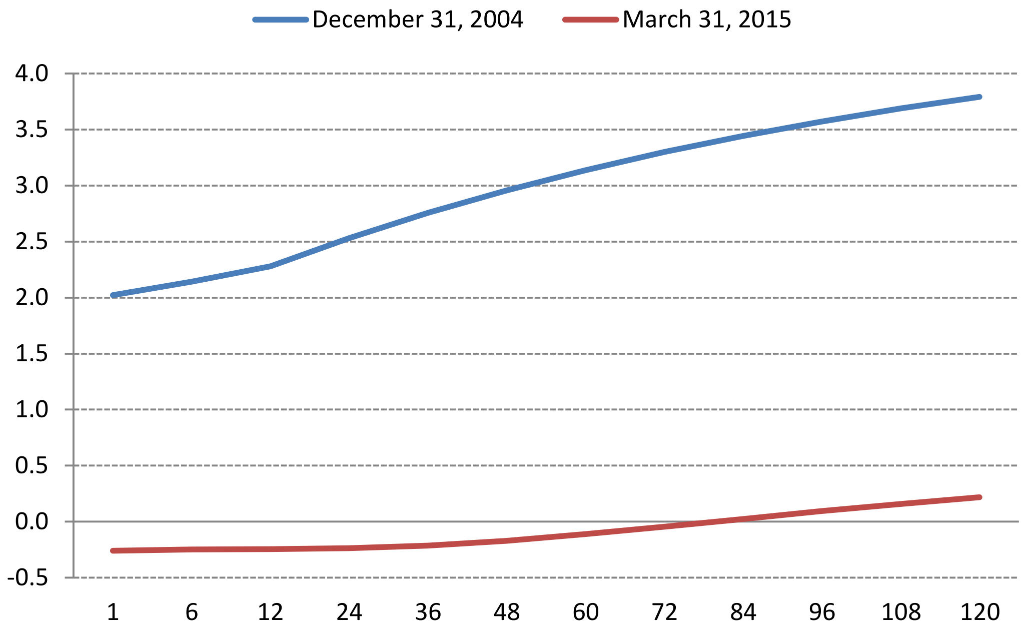

Figure 6: German zero-coupon yield curve in December 31, 2004 compared to the yield curve in March 31, 2015 for maturities from 1 month up to 120 months.

The negative excess return expectations given by the model between 2009 and 2015 are reasonable based on historical observations. As the model is calibrated initially with the pricing factors for such a different time period, it expects the rates to converge back to more normal levels from historical perspective. The model slowly adapts to this new environment and it does not expect that we would return anytime soon (if ever) to a completely similar yield environment as before the Eurozone crisis. This is indicated by extremely depressed levels of expected future short rates (see section 4.5 for more).

4.5. Convexity Bias

We estimate the effect of convexity for German yields based on the sample between August 29, 1997 and March 31, 2015 using the selected model by Adrian et al. (2013). The convexity bias is the second component in Equation (19):

$$-\frac{1}{2}\left(\beta^{(n-1)'}\Sigma\beta^{(n-1)}+\sigma^2\right). \quad (50)$$

Now we calculate the estimated convexity bias using the German data between August 29, 1997 and March 31, 2015. The results are shown in Figure 7. As expected the convexity bias is almost non-existent for maturities less than 3 years after which it slowly starts to slowly increase. With 10 year bonds the effect is already roughly -0.22 %. This means that without the convexity effect yields would be 0.22 % higher for ten year bonds just based on expected future average short rates and the risk premium.

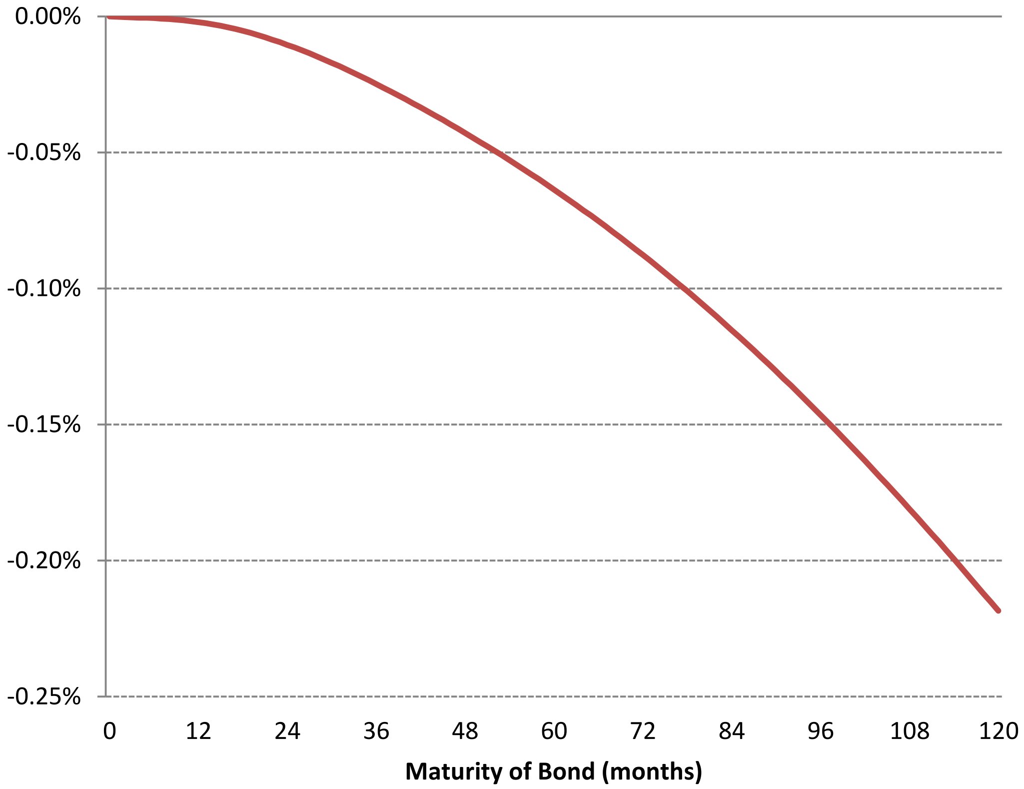

Figure 7: Convexity bias according to model by Adrian et al. (2013) for German yields when estimated for period between August 29, 1997 and March 31, 2015.

4.6. Analysis of the risk premium over the Eurozone crisis

In this section we proceed to analyze the estimated risk premium using the model by Adrian et al. (2013). In other words how bond pricing has changed during the Eurozone crisis compared to history. Figure 8, Figure 9 and Figure 10 show the model estimated risk premiums for the whole period 1997 – 2015 in 3 year, 5 year and 10 year maturities. Looking first at the whole period gives a good overview on how different the past 6 years really have been compared to time before the Eurozone crisis.

Looking at the short end maturities, we can see that German yields have fluctuated considerably since 1997. 3 year yield peaked at the aftermath of IT bubble and started declining until second half of 2005. The booming period drove yields higher until Lehman Brothers' bankruptcy late 2008. The financial crisis ignited by subprime lending and the following aggressive central bank policies drove yields quickly significantly lower also in Germany.

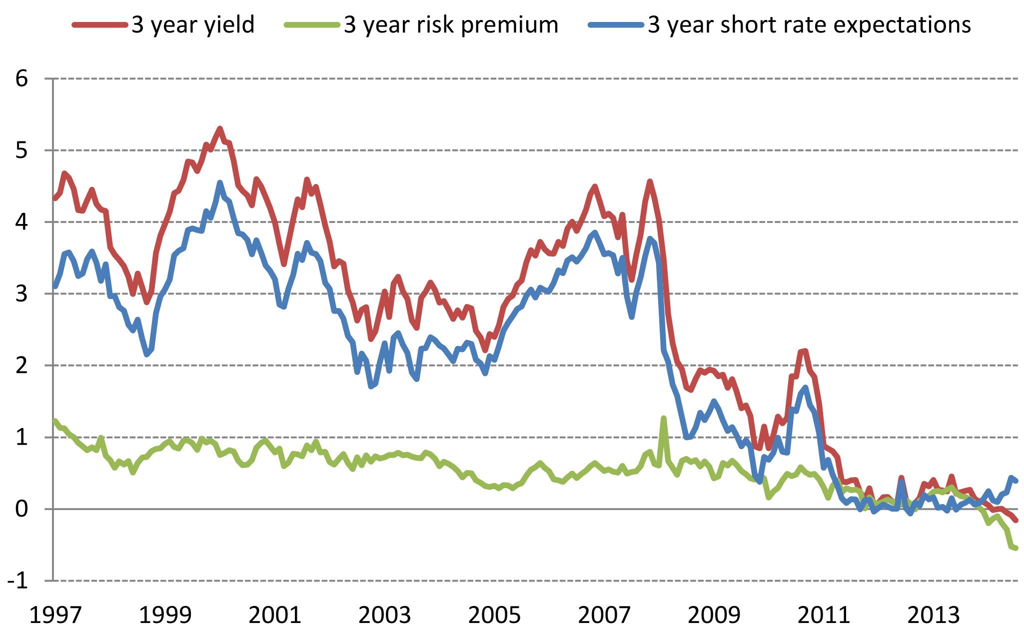

Figure 8: Yield decomposition 1997 - 2015 (n = 36 months)

Same pattern is repeated with 5 year bonds as can be seen from Figure 9. The risk premium had been rather stable all the way until 2009 excluding the brief spike caused by Lehman bankruptcy. Therefore the yields for 3 and 5 year bonds were mainly driven by short rate expectations.

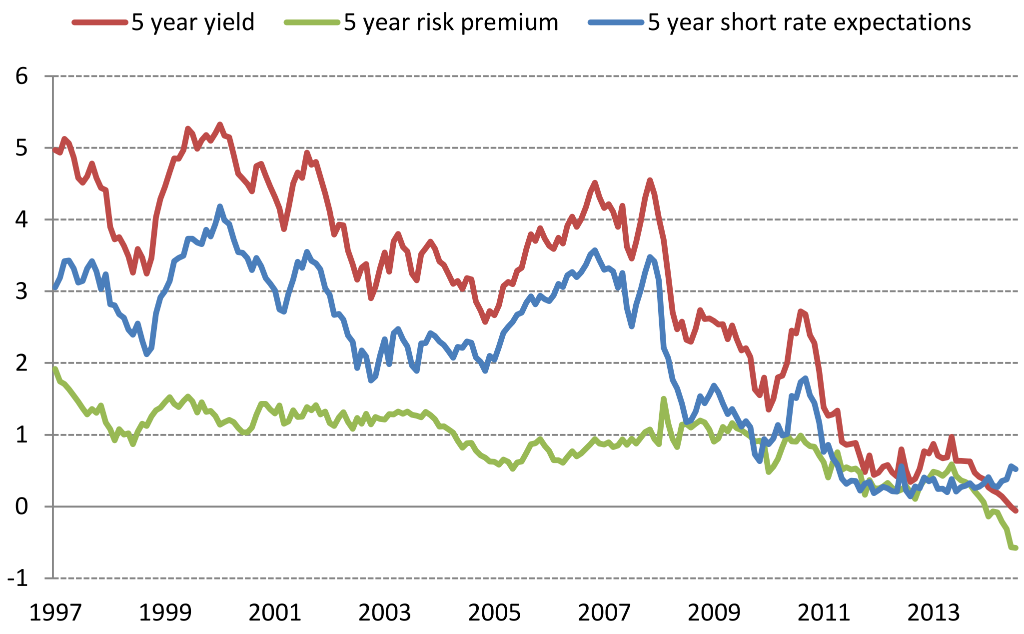

Figure 9: Yield decomposition 1997 - 2015 (n = 60 months)

Yields of the 10 year bonds have been more stable before the Eurozone crisis as shown in Figure 10. Yields had a small spike after the IT bubble and declined slightly before the second half of 2005. Short rate expectations and risk premiums were rather stable as well. The ten year risk premium was averaging roughly 2 % until 2005 after which the level shifted to roughly 1 % up until recent years. It made a brief attempt to climb back to the level of 2 % right before the start of Eurozone crisis. Short rate expectations were rising between 2005 and 2008.

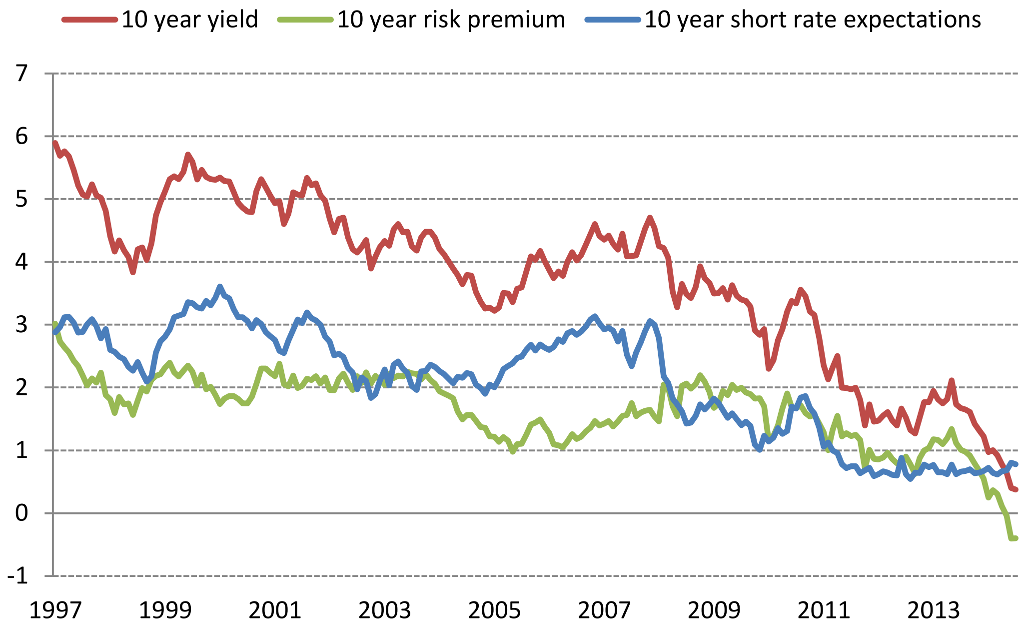

Figure 10: Yield decomposition 1997 - 2015 (n = 120 months)

The Eurozone crisis started from Greece around late 2009 and expanded to Ireland and Portugal and raised also concerns about Spain and Italy as well and the whole European banking system. These crisis countries had difficulties in financing their debt and thus suffered from rising interest rates and became dependent on outside help. (Belkin, et al., 2012).

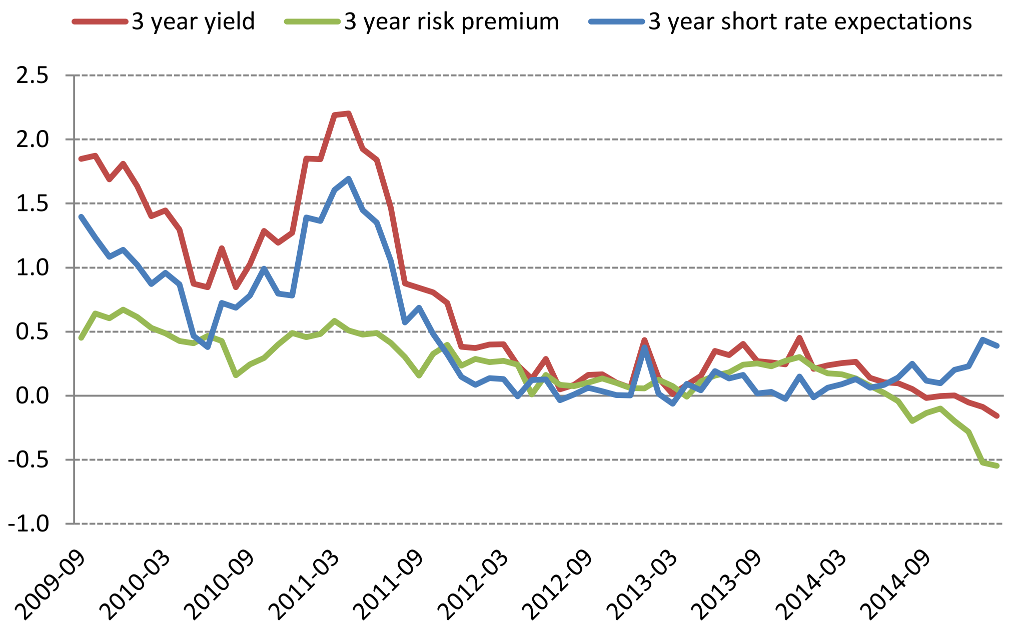

Figure 11: Yield decomposition September 2009 - March 2015 (n = 36 months)

We can see from Figure 11 how the 3 year bond yield and estimated risk premium as well as the average short rate expectations for three years changed during Eurozone crisis. Yields were coming down steadily as the Greece trouble emerged in 2009 up until June 2010. During April 2010 it became clear that Greece cannot serve its foreign debt without outside help (Bloomberg L.P., 2010). After this the 3 year bond yields raised up until March 2011. This rise was caused by expected increase in the future path of short rates maybe indicating a short relief on the crisis. The risk premium stayed rather stable up until March 2011. However this rise in 3 year yields was sharply reversed and by the end of 2011 the 3 year yield had dropped below 0.5 %. According to the estimated risk premium this was mainly caused by sudden drop in expected short rates up to three years. This possibly indicated the severity of the Eurozone crisis and the need for low rates and easing monetary policy actions for years to come. The estimated risk premium continued to decline in the short end and yields continued down all the way until these days. This happened even though the immediate risks in Eurozone crisis have declined for most countries and the yields of debt crisis countries have been more stable. The tipping point was the legendary statement by European Central Bank president Mario Draghi in July 2012 that he will do “whatever it takes” to preserve euro (Bloomberg L.P., 2014). Interestingly though this year the estimated future short rate expectations for the next three years have been on the rise while the risk premium has continued sharply further to the negative territory.

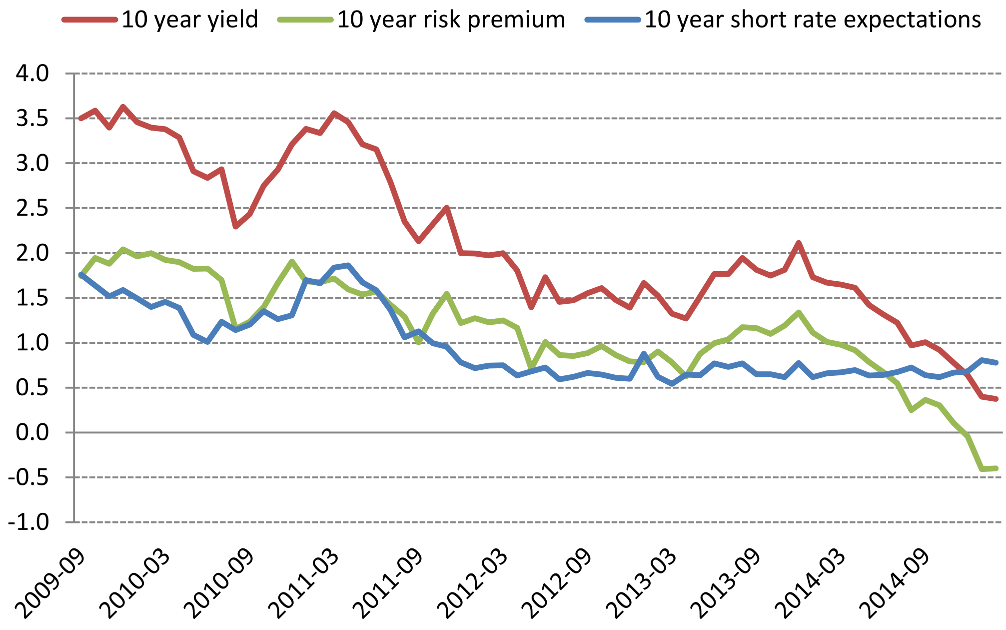

Figure 12: Yield decomposition September 2009 - March 2015 (n = 120 months)

Figure 12 shows the estimated risk premium and short rate expectations up to ten years during the Eurozone crisis. As can be seen the drop in ten year yields has been exceptional: from 3.5 % down to less than 0.5 %. The risk premium slowly declined as the Eurozone crisis progressed and continued to decline all the way to early 2014. After that it accelerated the decline and quickly went down to almost -0.5 %. This indicates that the stretched decline in 10 year bonds is mainly caused by risk premium while short rate expectations have been very stable since 2011. This huge decline in risk premium after early 2014 could have been driven by the expectations that the European Central Bank starts its own quantitative easing program that it finally announced in January 2015 (European Central Bank, 2015). In other words investors accepted negative expected excess returns with the anticipation of ECB starting to buy the bonds that would further drive yields even lower.

5. Conclusions

We chose the model by Adrian et al. (2013) to model the German risk premium. The main reasons are that it has shown good results with U.S. data and it is computationally fast to estimate even though it uses affine model framework. The model provides excellent fit compared to the used German data both when measured against fitted excess returns and fitted yields.

Out-of-sample forecasting during Eurozone crisis gave a bit mixed results. The model forecasted short term rates well compared to the two benchmark models especially at shorter maturities. However on average the model expected negative one month excess returns for long term bonds during the Eurozone crisis. In reality the yields of long term bonds have continued further down and the bonds have been earning high excess returns. Likely explanation is that the last six years have been very different compared to the period between 1997 and 2009 for which the model was initially estimated. Still the results compared to the benchmark models were not significantly different.

The estimated convexity bias from the German data is small compared to the risk premium except for the very long maturities. According to the analyzed risk premium the yields were driven mainly by short rate expectations during the first years of Eurozone crisis. The past two years have been completely different as the risk premium has collapsed pushing yields further down even though the short rate expectations have started to rise a little.

Further research could be done by including macroeconomic factors when modeling the German term structure with the model by Adrian et al. (2013). Also adding global pricing factors for example from U.S. data is an interesting idea. It can be argued that is the period between 1997 and 2015 sufficient to explain today's yield dynamics for German government bonds. Rolling estimation window approach might be better approach.

6. References

Adrian, T., Crump, R. & Moench, E., 2013. Pricing the Term Structure with Linear Regressions. *Journal of Financial Economics.*

Anderson, T., 1951. Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions.. *The Annals of Mathematical Statistics.*

Ang, A. & Piazzesi, M., 2003. A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. *Journal of Monetary Economics 50.*

Bauer, M. & Rudebusch, G., 2013. Monetary Policy Expectations at the Zero Lower Bound. *Federal Reserve Bank of San Francisco Working Paper.*

Bekaert, G., Hodrick, R. & Marshall, D., 1997. On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates. *Journal of Financial Economics 44.*

Bekaert, G., Wei, M. & Xing, Y., 2007. Uncovered Interest Rate Parity and the Term Structure. *Journal of International Money and Finance 26.*

Belkin, P., Weiss, M. A., Nelson, R. M. & Mix, D. E., 2012. The Eurozone Crisis: Overview and Issues for Congress. *Congressional Research Service Report R42377*.

Bloomberg L.P., 2010. *Greek Bailout Talks Could Take Three Weeks as Bond Repayment Looms in May.* [Online] Available at: http://www.bloomberg.com/news/2010-04-20/greek-talks-on-61-billion-aid-may-take-three-weeks-as-bond-payment-looms.html [Accessed 12 May 2015].

Bloomberg L.P., 2014. *Draghi's 'Whatever It Takes' Still Works as Euro Revives.* [Online] Available at: http://www.bloomberg.com/news/articles/2014-01-10/draghi-s-whatever-it-takes-still-works-as-euro-revives [Accessed 14 May 2015].

Boos, D., 2011. *Forecasting Asset Returns in State Space Models, Ph.D. thesis,* s.l.: s.n.

Brandt, M. & Chapman, D., 2008. Affine Term Structure Models. In: *The New Palgrave Dictionary of Economics.* s.l.:Palgrave Macmillan.

Bundesbank, 2015. *Prices and Yields of Listed Federal Securities.* [Online] Available at: http://www.bundesbank.de/Navigation/EN/Service/Federal_securities/Prices_and_yields/prices_and_yields.html

Campbell, J., 1986. A Defense of Traditional Hypotheses about the Term Structure of Interest Rates. *Journal of Finance 41.*

Campbell, J. & Shiller, R., 1991. Yield Spreads and Interest Rate Movements: a Bird's Eye View. *Review of Economic Studies 58.*

Christensen, J. & Rudebusch, G., 2013. *Estimating Shadow-Rate Term Structure Models with Near-Zero Yields.* s.l.:s.n.

Cochrane, J. H., 2000. *Asset Pricing.* s.l.:s.n.

Cochrane, J. & Piazzesi, M., 2005. Bond Risk Premia. *The American Economic Review 95.*

Cochrane, J. & Piazzesi, M., 2008. *Decomposing the Yield Curve.* s.l.:s.n.

Cox, J., Ingersoll, J. & Ross, S., 1985. A Theory of the Term Structure of Interest Rates. *Econometrica 53.*

Dahlquist, M. & Hasseltoft, H., 2013. International Bond Risk Premia. *Journal of International Economics 90.*

Duffee, G., 2002. Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance 57.*

Duffee, G., 2011. Information in (and not in) the Term Structure. *Review of Financial Studies 24.*

Duffie, D. & Kan, R., 1996. A Yield-Factor Model of Interest Rates. *Mathematical Finance 6.*

European Central Bank, 2015. *ECB Announces Expanded Asset Purchase Program.* [Online] Available at: https://www.ecb.europa.eu/press/pr/date/2015/html/pr150122_1.en.html [Accessed 16 May 2015].

Fama, E., 1990. Term-structure Forecasts of Interest Rates, Inflation and Real Returns. *Journal of Monetary Economics 25.*

Fama, E. & Bliss, R., 1987. The Information in Long-Maturity Forward Rates. *The American Economic Review.*

Feldstein, M. S., 2015. *Ending the Euro Crisis? No. w20862,* s.l.: s.n.

Halberstadt, A. & Stapf, J., 2012. *An Affine Multifactor Model with Macro Factors for the German Term Structure: Changing Results during the Recent Crises,* s.l.: Deutsche Bundesbank.

Hamilton, J. & Wu, J., 2012. *Identification and Estimation of Gaussian Affine Term Structure Models,* s.l.: s.n.

Hellerstein, R., 2011. *Global Bond Risk Premiums,* s.l.: Federal Reserve Bank of New York.

Ichiue, H. & Ueno, Y., 2013. *Estimating Term Premia at the Zero Bound: An Analysis of Japanese, US, and UK Yields.* s.l.:Bank of Japan.

Ilmanen, A., 1995. Overview of Forward Rate Analysis. *Understanding the Yield Curve Part 1.*

Ilmanen, A., 1995. Time-Varying Expected Returns in International Bond Markets. *The Journal of Finance 50.*

Ilmanen, A., 1996. Market Rate Expectations and Forward Rates. *The Journal of Fixed Income 6.2,* pp. 8-22.

Ilmanen, A., 1997. Forecasting US Bond Returns. *The Journal of Fixed Income 7.1,* pp. 22-37.

Ilmanen, A., 2000. Convexity Bias and the Yield Curve. In: *Advanced Fixed-Income Valuation Tools 61.* s.l.:s.n., p. 25.

Jolliffe, I., 2002. *Principal Component Analysis.* s.l.:John Wiley & Sons.

Joslin, S., Singleton, K. J. & Zhu, H., 2011. A New Perspective on Gaussian Dynamic Term Structure Models. *Review of Financial Studies 24.3.*

Kessler, S. & Scherer, B., 2009. Varying Risk Premia in International Bond Markets. *Journal of Banking & Finance 33.*

Kim, D. & Orphanides, A., 2012. Term Structure Estimation with Survey Data on Interest Rate Forecasts. *Journal of Financial and Quantitative Analysis.*

Kim, D. & Singleton, K., 2012. Term Structure Models and the Zero Bound: an Empirical Investigation of Japanese Yields. *Journal of Econometrics 170.*

Kim, D. & Wright, J., 2005. An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates. *Finance and Economics Discussion Series 2005-33.*

Le, A., Singleton, K. J. & Dai, Q., 2010. Discrete-Time Affine Term Structure Models with Generalized Market Prices of Risk. *Review of Financial Studies 23.5.*

Litterman, R. & Scheinkman, J., 1991. Common Factors Affecting Bond Returns. *Journal of Fixed Income 1.*

Nelson, C. & Siegel, A., 1987. Parsimonious Modeling of Yield Curves. *The Journal of Business, Vol. 60, No. 4..*

Piazzesi, M., 2010. Affine Term Structure Models. In: *Handbook of Financial Econometrics.* s.l.:Elsevier.

Radwanski, J., 2012. *An Economic Interpretation of Bond Return Predictability,* s.l.: s.n.

Richter, A. W. & Throckmorton, N. A., 2015. The Zero Lower Bound Frequency, Duration and Numerical Convergence. *The BE Journal of Macroeconomics 15.1.*

Schich, S., 1997. Estimating the German Term Structure. *Economic Research Group of the Deutsche Bundesbank.*

Svensson, L., 1994. Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994. *National Bureau of Economic Research.*

Tuckman, B., 2002. *Fixed income securities: tools for today's markets..* s.l.:John Wiley & Sons.

Vasicek, O., 1977. An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics 5.*

Verbeek, M., 2004. *A Guide to Modern Econometrics.* s.l.:s.n.

Wäger, L., 2012. *Analysis of Bond Risk Premia: Extensions to Macro-Finance and Multi-Currency Models, Ph.D. thesis.* s.l.:University of St. Gallen.